

POSUDEK VEDOUCÍHO BAKALÁŘSKÉ PRÁCE

Název: Problém vícenásobného testování

Autor: Oliver Waczulík

SHRNUTÍ OBSAHU PRÁCE

Práce pojednává o problematice vícenásobného testování, se kterým se setkáváme vždy, když potřebujeme provést najednou (simultánně) statistické testy o platnosti několika nulových hypotéz. Uvedení do problematiky a zejména zavedení několika v praxi často používaných měr výskytu chyby prvního druhu v kontextu vícenásobného testování (v podstatě zobecnění pravděpodobnosti chyby prvního druhu) vůči nimž jsou následně konstruovány statistické testy, je obsaženo v 1. kapitole předložené práce. Druhá kapitola následně stručně popisuje nejjednodušší (avšak v praxi patrně nejčastěji používaný) přístup založený na Bonferroniho nerovnosti. Ve třetí kapitole se autor věnuje tzv. krokovým procedurám a zejména postupu publikovanému Holmem v roce 1979 (do seznamu referencí se autorovi vloudila chyba, kdy je chybně uveden rok 1972) a v následující čtvrté kapitole potom tzv. principu uzavírání. Poslední, pátá kapitola je věnována specifikům vícenásobného testování v jedné z klasických situací, kterou je analýza rozptylu, tj. testování shody středních hodnot několika normálně rozdělených náhodných výběrů.

CELKOVÉ HODNOCENÍ PRÁCE

Téma práce. Téma práce vychází ze základních pojmů používaných v kontextu testování statistických hypotéz, na jejichž základě autor provádí rozšíření nad rámec znalostí běžně vyučovaných na bakalářském studiu oboru Finanční matematika. Získané poznatky jsou však poměrně široce využitelné v kontextu statistické analýzy dat, se kterou se absolvent zmíněného oboru zcela jistě bude v nemalé míře potkávat. Téma se tedy jeví jako vhodné pro posluchačův obor studia. S ohledem na zadání práce byl nesplněn pouze bod týkající se provedení simulační studie s cílem prozkoumat, jak přesně dodržují jednotlivé procedury předepsanou mez výskytu chyb prvního druhu. K nezahrnutí této simulační studie však došlo po vzájemné dohodě s vedoucím práce v okamžiku, kdy již bylo zřejmé, že i bez tohoto bodu bude předložená práce patřit svým rozsahem spíše k těm delším (konečný rozsah je 38 stran bez strany titulní a dalších povinných stran na začátku práce).

Vlastní příspěvek. Hlavním autorovým příspěvkem je utřídění a formalizace (při jednotném značení) poznatků ze studované oblasti a dále též mnohem podrobnější uvedení některých matematických odvození dostupných v literatuře pouze heslovitou formou (viz též oddíl „Práce se zdroji“ níže). Kromě toho autor prakticky aplikoval studované postupy na reálná data, která si sám zajistil.

Matematická úroveň. Matematická úroveň textu je dobrá. Používané (méně standardní) pojmy jsou řádně definovány, z textu je vesměs patrné, co se předpokládá a co se následně tvrdí. Jednotlivá tvrzení jsou vesměs dokázána způsobem, který by měl být srozumitelný též pro běžné studenty autora oboru. V práci se dle mého názoru nevyskytují žádné zásadní matematické chyby.

Práce se zdroji. Používané zdroje jsou řádně a jednotným způsobem stylem Autor (Rok) u přímých a (Autor, Rok) u nepřímých referencí citovány. Celkově se jedná o práci kompilační, kdy se autorovi podařilo (dle mého názoru) hezkým způsobem utřídřit poznatky ze studované problematiky. Hlavní zdroje, ze kterých autor vycházel, byly čtyři: kniha Bretz, Hothorn a Westfall (2011).

Zde autor čerpal zejména základní názvosloví a popis jednotlivých procedur pro svoje kapitoly 1 – 4. Kniha je nicméně napsána poměrně popisným způsobem, který se autor předložené bakalářské práce vesměs úspěšně snažil zformalizovat a pokusil se doplnit rigoróznější matematická odvození. Pro kapitolu 4 čerpal autor nemálo též z původního Holmova článku z roku 1979. Oddíl 5.2 vychází opět z původního článku Dunnett (1955), zbylé části kapitoly 5 potom zejména z klasické Andělovy učebnice (Anděl, 2007). V rámci bakalářské práce se autor snažil podat vše v jednotné méně kompaktní formě a dále též ve formě srozumitelné pro studenty s úrovní statistických znalostí standardních u autorova oboru studia. Zejména v případě časopiseckých (34, resp. 58 let starých) zdrojů, ale též v případě knihy Bretz, Hothorn a Westfall se jednalo o nelehký úkol, se kterým se dle mého názoru autor úspěšně vyrovnal.

Formální úprava. Po formální stránce je práce na velice dobré úrovni. V souladu se zadáním byla připravena pomocí systému $\text{L}^{\text{A}}\text{T}_{\text{E}}\text{X}$, se kterým se student musel samostatně (a soudě dle výsledku úspěšně) seznámit při začátku psaní práce. S ohledem na jazyk práce (slovenština) mohu pouze v omezené míře posoudit pravopis, nicméně počet zjevných tiskových chyb je přiměřený. Kromě výše zmíněné chyby v roce vydání Holmova článku jsem při čtení závěrečné práce našel ještě nesprávně uvedené jméno Torstena Hothorna v seznamu referencí, kde je uveden jako „Horthon“ a dále malé množství překlepů, které lze s ohledem na rozsah a povahu práce jistě prominout.

ZÁVĚR

Závěrem je nutno podotknout, že student pracoval ve velké míře samostatně s pouze mírnými vstupy ze strany vedoucího. Práci považuji za velmi dobrou a **doporučuji ji uznat** jako bakalářskou práci.

doc. RNDr. Arnošt Komárek, Ph.D.

Katedra pravděpodobnosti a matematické statistiky
Matematicko-fyzikální fakulta Univerzity Karlovy v Praze

V Praze dne 24. srpna 2013